Real-time ventilation control for indoor CO_2 management using occupant information

Kang Woo Bae [a], Eun Ji Choi [a], Young Jae Choi [a], Ji Young Yun [a], Geun Young Yun [b], Hyeun Jun Moon [c], Jin Woo Moon [d,*]

[a] School of Architecture and Building Science, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, Seoul, 06974, South Korea
[b] Department of Architectural Engineering, Kyung Hee University, 1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, South Korea
[c] Department of Architectural Engineering, Dankook University, Yongin 16890, South Korea
[d] School of Architecture and Building Science, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, Seoul 06974, South Korea

ARTICLE INFO

ABSTRACT

In contemporary Conventional ventilation control methods primarily rely on environmental measurements taken indoors but often overlook occupant-specific factors. This study presents an optimal ventilation control algorithm based on a real-time indoor CO_2 concentration prediction model designed to enhance IAQ and energy efficiency. This model incorporates real-time occupant data, highlighting the considerable influence of occupant-related variables on the indoor CO_2 levels. To this end, three deep learning architectures—deep neural networks, long short-term memory (LSTM), and gated recurrent units—were evaluated, with the LSTM model exhibiting superior accuracy and robustness. Using this model, a predictive ventilation control algorithm was developed to proactively regulate airflow and maintain CO_2 concentrations below the recommended threshold of 1,000 ppm. The effectiveness of the proposed control strategy was validated using mockup experiments and living lab-based simulations. The results show that integrating real-time occupant data considerably enhances indoor comfort than rule-based ventilation control. Furthermore, optimal ventilation control resulted in a considerable decrease in energy consumption by approximately 24.74%, particularly in large-scale environments. These findings highlight the potential of the proposed method as a robust solution for next-generation indoor environmental management systems and intelligent control in smart buildings.

1. Introduction

As modern lifestyles have shifted predominantly indoors, maintaining optimal indoor air quality (IAQ) has become crucial for occupant health and comfort. With people spending approximately 90% of their time in enclosed spaces, effective IAQ management is crucial for health protection and building sustainability [1]. In response, numerous countries have strengthened their IAQ standards and implemented policy measures to ensure compliance [2]. These policies protect occupant health and support sustainable building operation. Key contributors to IAQ degradation include carbon dioxide (CO_2) buildup from occupant respiration, volatile organic compounds emitted from building materials, and particulate matter from external sources [3]. Prolonged exposure to these pollutants can result in health issues such as fatigue, headache, and decreased concentration. Additionally, CO_2 concentrations exceeding certain levels can induce dizziness and drowsiness, thereby affecting occupant health [4].

Indoor CO_2 concentrations are affected not only by environmental factors such as the ventilation rate and outdoor air quality, but also by personal factors, including the body mass index (BMI), metabolic rate (MET), and gender [5]. Studies show that a 10% increase in body weight results in an average 8% increase in CO_2 emissions, indicating a correlation between a higher BMI and increased respiration and CO_2 output [6]. Park et al [7]. observed that CO_2 emissions increase by approximately 3-8 times in high MET activities such as walking (3.5 MET) and running (8.0 MET) than sitting. Owing to greater lung capacity and muscle mass, men exhale more CO_2 than women [8]. Hence, incorporating occupant-specific information is crucial to control personalized ventilation systems.

Conventional ventilation system control methods—on/off, demand-controlled ventilation (DCV), and proportional–integral–derivative

---

Nomenclature

$\dot{m}$ opt	Optimal supply airflow rate
$\dot{m}$ SA	Supply airflow rate
$\dot{m}$ RA	Return airflow rate
CO_2 pred	Predicted indoor CO_2 concentration
CO_2 limit	Predefined CO_2 concentration threshold (generally 1,000 ppm)
C	Performance load factor coefficients used in energy consumption calculations
FF	Airflow ratio (actual airflow to maximum possible airflow)

Abbreviation

IAQ	Indoor air quality
CO_2	Carbon dioxide [ppm]
LSTM	Long short-term memory
GRU	Gated recurrent unit
DNN	Deep neural network
MET	Metabolic rate [met or W/m^2]

PID	Proportional–integral–derivative
DCV	Demand-controlled ventilation
MPC	Model predictive control
ERV	Energy recovery ventilation
HVAC	Heating, ventilation, and air conditioning
BMI	Body mass index [kg/m^2]
AI	Artificial intelligence
M^3/H	Cubic meters per hour [m^3/h]
MAE	Mean absolute error
MAPE	Mean absolute percentage error
CvRMSE	Coefficient of variation of the root mean square error
R^2	Coefficient of determination
FPS	Frames per second
SA	Supply airflow [m^3/h]
RA	Return airflow [m^3/h]
EHP	Electric heat pump
PLF	Performance load factor
FF	Airflow ratio

(PID) control—are widely used to regulate indoor airflow and maintain acceptable air quality levels. On/off control is simple and cost-effective but can result in energy waste owing to rapid CO_2 concentrations fluctuations. DCV monitors IAQ data in real time and adjusts the ventilation rates based on demand, thereby offering improved energy efficiency. PID control provides stable regulation by minimizing errors through real-time feedback. However, these conventional control methods rely primarily on environmental variables measured indoors and fail to adequately incorporate occupant-specific information.

To address the aforementioned limitation, approaches such as model predictive control (MPC) and artificial intelligence (AI) have been introduced. MPC incorporates occupant information and both indoor and outdoor environmental data to predict CO_2 concentrations and proactively adjust ventilation rates, thereby enhancing the air quality and energy efficiency [9,10]. However, it is computationally demanding owing to the Consequently, AI-based smart ventilation control technologies have emerged, leveraging real-time data collection and analysis to deliver optimized ventilation strategies [11].

The occupant-centric control (OCC) concept, which utilizes real-time occupant data to optimize the building energy consumption and indoor environmental quality, has attracted considerable attention [12]. Numerous studies have applied the OCC methods that adjust ventilation rates based on occupant CO_2 emissions, occupancy status, and density [13–15]. Choi et al [16]. installed video-based occupant sensors in a living lab and used a convolutional neural network to estimate occupant numbers. By controlling an energy recovery ventilation (ERV) system based on this data, they achieved approximately 40% energy savings than PID control. Kim et al [17]. developed a double deep Q-network algorithm that identifies simple occupancy states (work, sleep, and absence) using a multivariate logistic regression model. This algorithm simultaneously controls cooling, ventilation, and humidification systems, thereby resulting in energy savings ranging from 6.3-21% [17]. Mutis et al [18]. proposed a heating, ventilation, and air-conditioning (HVAC) control model that combines YOLO v3 with a multistream deep learning network (Conv2D, long short-term memory (LSTM), and Conv3D) for occupant activity detection and prediction, thereby achieving energy savings and improved occupant comfort. These studies highlight the importance of incorporating different types of occupant information (occupant count, presence/absence, MET, gender, and BMI) for accurate CO_2 control and IAQ enhancement.

Existing studies primarily focus on detecting the current state and activity of occupants, without fully considering how these activities affect indoor CO_2 concentration levels [19–21]. Consequently, ventilation systems generally operate reactively, acting only after the IAQ thresholds are exceeded, making difficult to maintain comfortable indoor environments. To address this issue, further research is required to predict changes in indoor CO_2 concentrations by incorporating occupant information, thereby enabling the proactive control of the ventilation system to enhance IAQ and comfort.

This study aimed to develop an intelligent model for real-time CO_2 concentration prediction based on occupant information and propose an optimized ventilation control strategy. It seeks to comprehensively analyze the impact of the proposed method on IAQ, comfort, and energy consumption using field applications. By integrating real-time CO_2 prediction with ventilation control, the study addresses the limitations of conventional reactive control methods and demonstrates the effectiveness of a proactive control model for maintaining IAQ and comfort.

Fig. 1 shows the process of developing the prediction model and optimized ventilation algorithm (algorithm-opt), including a performance evaluation. In Step 1, data collection and preprocessing were conducted to gather indoor and outdoor environmental data, occupant information (number, MET, gender, and BMI), and ventilation system data. Preprocessing improved data quality, thereby enabling the development of the training dataset and analysis of correlations among key variables. Step 2 involved developing the prediction model using different deep learning techniques such as deep neural networks (DNNs), gated recurrent units (GRUs), and LSTM networks to optimize the CO_2 concentration prediction model based on the collected data. In Step 3, the optimal control algorithm was developed to calculate optimal ventilation rates using real-time predicted data, thereby enabling proactive ventilation control. Step 4 comprised a performance evaluation using mockup and living lab tests wherein the impact of the algorithm on IAQ improvement and energy consumption were analyzed. The mock-up test analyzed CO_2 concentrations and energy consumption in a single-user environment, while the living lab test evaluated the scalability and stability of the algorithm in a multi-user scenario using computer simulations.

Numerous studies have applied OCC strategies utilizing data such as occupant count, presence, and activity state. However, these approaches rely primarily on detecting the current state and reactively operating the ventilation system once the IAQ thresholds are exceeded. Reactive systems often result in delayed responses, thereby resulting in temporary drops in the IAQ and reduced indoor air comfort when CO_2 levels surpass the acceptable limits.

While there has been increasing interest in real-time AI-based ventilation control, existing studies have largely overlooked the
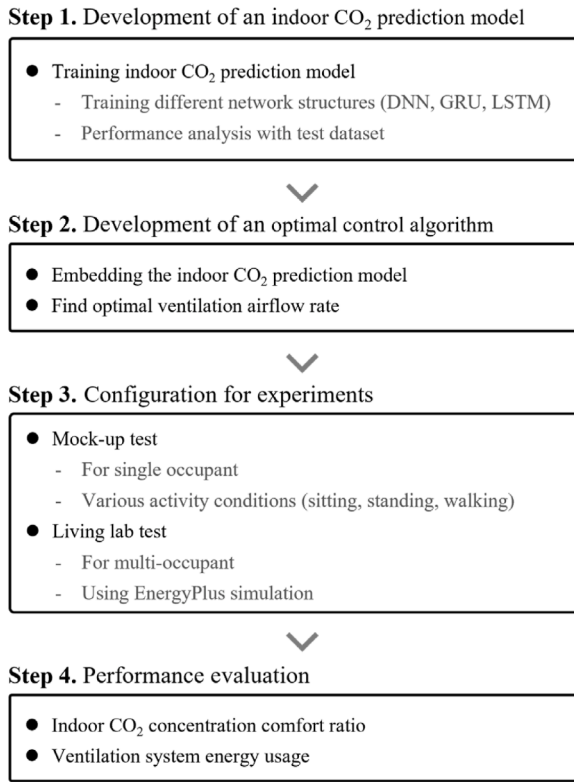
Step 1. Development of an indoor CO_2 prediction model

- Training indoor CO_2 prediction model
  - Training different network structures (DNN, GRU, LSTM)
  - Performance analysis with test dataset

∨

Step 2. Development of an optimal control algorithm

- Embedding the indoor CO_2 prediction model
- Find optimal ventilation airflow rate

∨

Step 3. Configuration for experiments

- Mock-up test
  - For single occupant
  - Various activity conditions (sitting, standing, walking)
- Living lab test
  - For multi-occupant
  - Using EnergyPlus simulation

∨

Step 4. Performance evaluation

- Indoor CO_2 concentration comfort ratio
- Ventilation system energy usage

Fig. 1. Process flow of the methodology.

dynamic integration of occupant-specific characteristics such as MET, BMI, and gender into CO_2 prediction models to enable proactive control strategies. The majority of prior studies have focused on reactive approaches, with limited investigation into personalized ventilation systems that incorporate real-time occupant behavior and physiological attributes.

This study addresses the aforementioned gap by proposing a method that optimizes ventilation control by integrating occupant-specific information (MET, BMI, and gender) into real-time CO_2 prediction models. By enabling the system to predict and manage IAQ degradation before it occurs, the approach supports consistent indoor air quality, enhances comfort, and optimizes energy efficiency. Unlike conventional reactive systems, this proactive model provides real-time, predictive IAQ management, thereby providing a novel ventilation control strategy that distinguishes it from previous studies.

The remainder of this paper is organized as follows. Section 2 outlines the methodology for developing the indoor CO_2 concentration prediction model. First, input and output data are selected using correlation analysis, followed by the development of the prediction model using deep learning algorithms (DNN, LSTM, and GRU). It describes the proposed optimized ventilation control algorithm (algorithm-opt) and rule-based control method (algorithm-rule) used for comparison.

Section 3 provides an overview of the mock-up and office (living lab) environments used to compare the control performance of the optimized ventilation algorithm (algorithm-opt) and rule-based control method (algorithm-rule). It outlines the participant details for the mock-up experiment and living lab computer simulation.

Section 4 presents the results of the ventilation control system performance evaluation based on experiments conducted in the mock-up and office (living lab) environment. The experiments measure CO_2 concentrations relative to occupant activity levels, comparing the performance of the two control methods (algorithm-rule and algorithm-opt). Herein, the findings are analyzed and suggestions for future research are proposed.

Section 5 summarizes Sections 2 to 4, evaluates the ventilation

control system performance using the experimental results and comparisons, analyzes the differences between the two control methods, and discusses potential future research directions. Based on the findings, it offers recommendations for advancing IAQ management and optimizing control systems.

2. Development of optimal control method

2.1. Indoor CO2 prediction model

Indoor CO_2 concentration exhibits dynamic, nonlinear characteristics that are influenced by numerous factors, requiring deep learning-based models capable of capturing this complexity. In this study, prediction models were developed using DNN [22], LSTM [23,24], and GRU [25,26]. While DNNs offer high learning capacities owing to their multilayered architecture, they are prone to overfitting and extended training times. To address this, regularization techniques and dropout were employed. LSTM networks effectively capture the temporal dependencies in historical data, although their complex structures result in longer training times. Conversely, GRUs offer high computational efficiency suitable for real-time data processing, offering a performance comparable to LSTM while addressing its limitations [27].

Appropriate input variable selection is crucial for achieving robust prediction performance. In this study, candidate input variables were identified based on environmental and personal factors affecting indoor CO_2 concentration. Pearson correlation analysis was utilized to select the most influential variables. The dataset for this analysis was collected using preliminary experiments conducted in a mock-up environment from February 2-19, 2024, resulting in 66,632 data entries. The data were obtained using environmental sensors and a vision-based occupancy estimation system installed in the mock-up environment, which served as the basis for the statistical analysis used in input selection.

The initial set of candidate variables included indoor temperature, outdoor temperature, indoor humidity, outdoor CO_2 concentration, supply airflow rate (SA), return airflow rate (RA), number of occupants, MET, BMI, and gender. As shown in Fig. 2 (a) and (b), the Pearson correlation coefficient ranges from –1 to 1, with values near –1 indicating a strong negative correlation and those near 1 indicating a strong positive correlation. Generally, coefficients above 0.8 indicate a strong correlation, between 0.4 and 0.8 a moderate correlation, and below 0.4 a weak correlation [28,29]. The analysis revealed that the number of occupants, MET, BMI, and gender had correlation coefficients exceeding 0.5, indicating considerable influence on indoor CO_2 concentration. Although SA and RA exhibited lower coefficients (~0.35), they were retained as input features owing to their indirect impact on air circulation and CO_2 levels.

In addition to the Pearson correlation analysis, Spearman rank correlation was conducted to identify monotonic nonlinear relationships between the input variables. The Spearman coefficient complements the Pearson correlation by capturing not only linear relationships but also order-based nonlinear trends [30]. Indoor CO_2 exhibited positive Spearman correlations above 0.4 with the number of occupants, MET, SA, and RA, indicating that the combined influence of occupancy and activity levels on indoor CO_2 concentration may follow nonlinear patterns.

The strong correlation between BMI and gender further indicates that the MET estimation method, that relies on physical characteristics of participants, considerably influences the results. This finding supports the reliability of the relationship between the input variables and target prediction, particularly in the context of posture-based MET estimation.

In the experimental system, the SA comprised only outdoor air (OA), with no mixing with RA, while the damper operated in binary mode (either 0% or 100%). The return air was completely exhausted outdoors and did not re-enter the supply stream, meaning no mixed air (MA) was formed. The final set of input variables comprised seven features: the indoor CO_2 concentration, SA, RA, number of occupants, MET, BMI, and
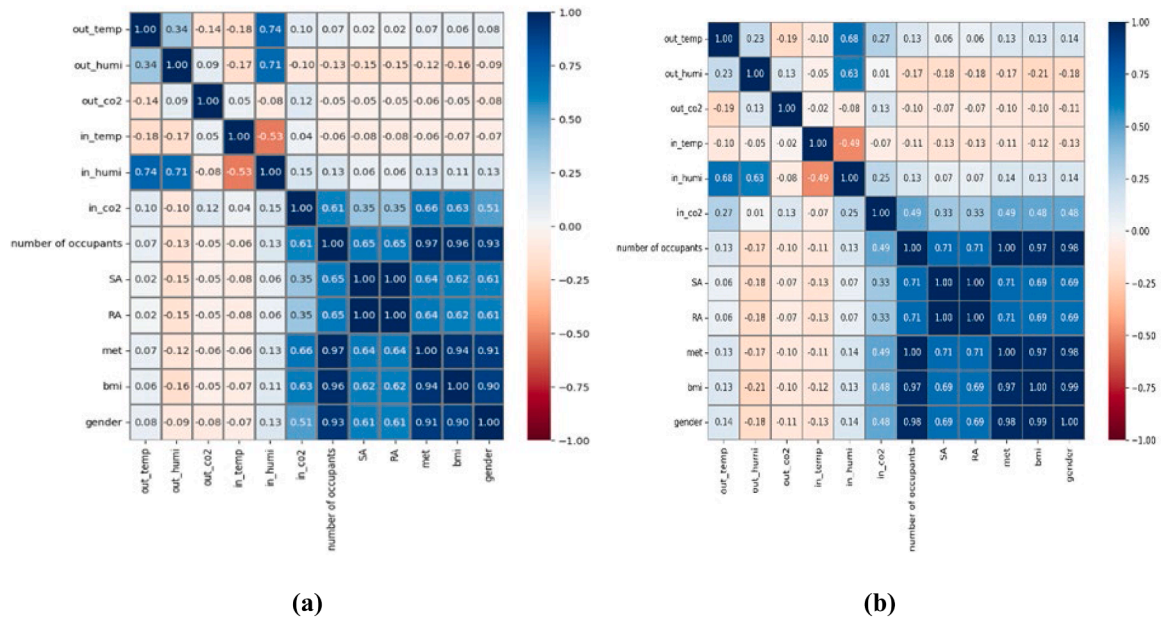
Fig. 2. (a) Pearson and (b) Spearman correlations of the input variables.

gender. The output variable was defined as the indoor CO_2 concentration 5 min ahead, which corresponds with the system control cycle. Using preliminary experiments, an appropriate control cycle was determined based on the airflow levels. Considering the volume and internal diffusion characteristics of the mock-up space, when the control cycle was set to more than 10 min, the system repeatedly failed to control the CO_2 concentration until it exceeded the indoor air quality threshold of 1,000 ppm. Consequently, a 5-min cycle was adopted as the minimum control unit to ensure stable indoor air quality management.

The number of occupants and MET values were automatically extracted using a vision-based sensor that analyzed real-time image data to estimate activity levels. For MET, the most frequent value within each 5-min interval was selected from measurements captured at a frame rate of 2 FPS [31]. Conversely, BMI and gender were manually input based on data collected from participants during preliminary experiments.

Outdoor CO_2 concentration exhibited negligible correlation and was excluded from the final input set. Although outdoor CO_2 is a key variable in diluting indoor CO_2 concentrations—particularly in spaces with high infiltration even without proper supply airflow—this study selected input features based on statistical significance using correlation analysis. Hence, only variables with a stronger quantitative influence were retained for model development.

The collected data were preprocessed for model training by replacing outliers—caused by communication instability—with the average of the preceding and following data points. No missing values were observed. Given a 5-min control cycle, the raw data were averaged over 5-min intervals, thereby resulting in 3,336 training data entries. The dataset was then denoised and structured to meet the input format requirements of the DNN, LSTM, and GRU models.

All input variables were normalized to a 0-1 range using min–max scaling. The dataset was then divided into training, validation, and test sets in an 8:1:1 ratio, which corresponds with 2,670, 333, and 333 entries, respectively. To maintain the temporal structure of the time-series data, a time-based sequential split was applied instead of random partitioning.

Before training the DNN, LSTM, and GRU models using the pre-processed data, a hyperparameter-tuning process was conducted to ensure optimal performance. Key hyperparameters, such as the number of neurons and hidden layers, dropout rate, and activation functions, are crucial for determining the structure and learning efficiency of a model [32]. As model performance can vary considerably based on

hyperparameter combinations, selecting an optimal set using an efficient search strategy is crucial [33]. Common hyperparameter optimization techniques include grid search, random search, heuristic search, and Bayesian optimization [34]. Herein, Bayesian optimization was adopted owing to its search efficiency and suitability under time constraints [35].

Bayesian optimization was conducted using the Bayesian Optimization module from the Keras Tuner library. The search space for the hypermodel was defined as follows: number of neurons ranging from 32 to 512, number of hidden layers from 0 to 5, dropout rate between 0 and 0.5, and activation functions such as ReLU, sigmoid, and tanh. The optimization process used the mean squared error (MSE) as the loss function and the Adam optimizer for training. The training configuration included 100 epochs, 15 repeated training runs, with early stopping triggered after five consecutive iterations without improvement. A fixed window size was used for the LSTM model.

Owing to the Bayesian optimization, the optimal configuration for the DNN model included three hidden layers, a dropout rate of 0.2, and the ReLU activation function. For the GRU model, regardless of the window size (2-4), the optimal configuration comprised two hidden layers, a dropout rate of 0.1, and the Tanh activation function. The LSTM model exhibited a similar configuration to the GRU. Further details are provided in Section 4.1.

2.2. Algorithm for ventilation system control

This section describes the development methodology of algorithm-opt, which optimally controls the ventilation system using the predicted indoor CO_2 concentration. Additionally, a rule-based ventilation control algorithm (algorithm-rule) was used as a baseline for the comparative evaluation of IAQ and energy performance in real-world environments.

The algorithm-opt integrates the CO_2 concentration prediction model described in Section 2.1 to enable ventilation system control. The flow of the algorithm is illustrated in Fig. 3. This approach is designed to maintain indoor CO_2 concentrations below the target threshold of 1000 ppm, in accordance with the acceptable indoor air quality limits recommended by ASHRAE Standard 62.1 [36]. To achieve this, the control variable (predicted indoor CO_2 concentration ($CO_{2\ pred\ i(t)}$)) was estimated. Based on this prediction, the manipulated variable (optimal
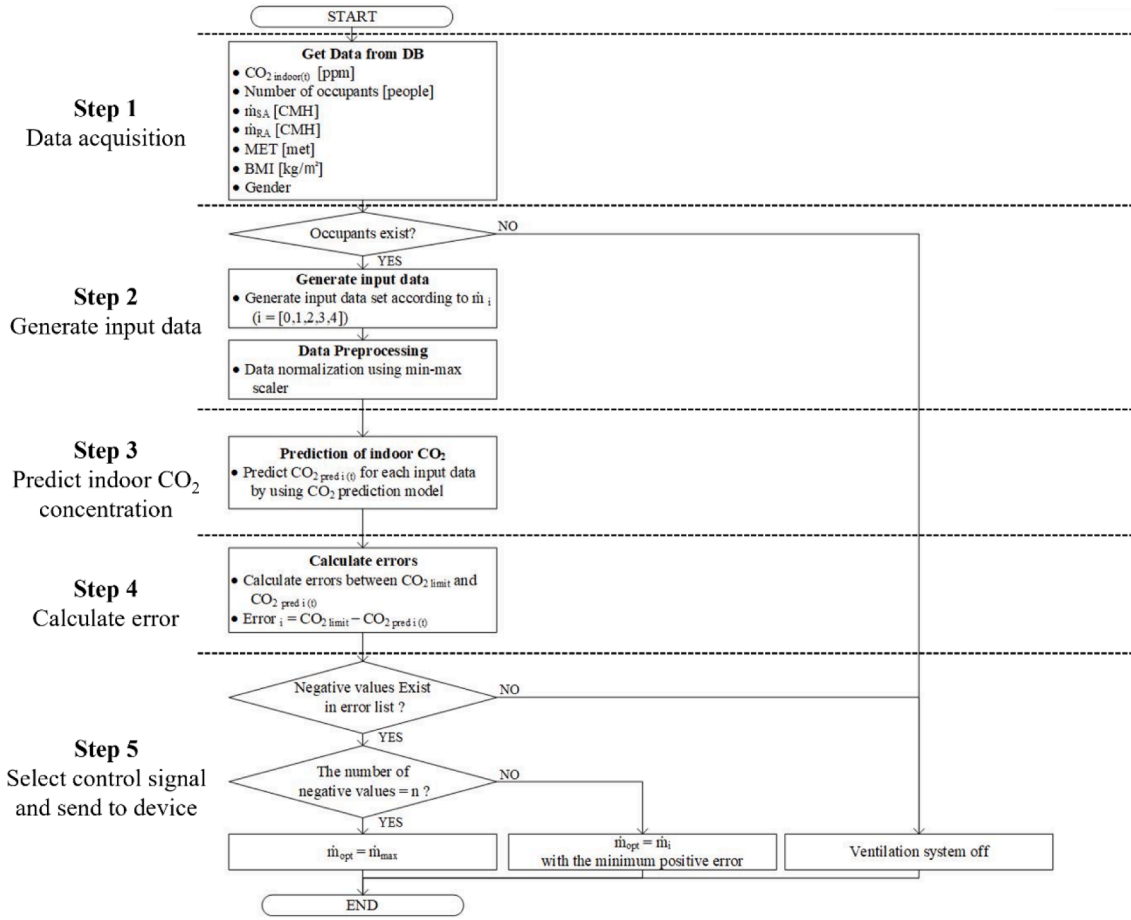
Fig. 3. Optimal ventilation control algorithm.

supply airflow rate ($\dot{m}_{opt}$)) was determined.

Step 1 involved retrieving real-time stored data from the database. The acquired data include the seven input variables identified in Section 2.1: indoor CO_2 concentration ($CO_{2\ indoor(t)}$), number of occupants, MET, supply and return airflow rates ($\dot{m}_{SA}$ and $\dot{m}_{RA}$), BMI, and gender. The indoor CO_2 concentration was represented as a 5-min average, while the number of occupants and MET were represented as the most frequent values within the same 5-min interval. The SA and RA values corresponded with the control settings of the ventilation system at the time of data acquisition. BMI and gender were fixed based on the registered occupant profiles.

Step 2 involved preprocessing the input data. First, occupant presence was assessed. If no occupants were detected, the ventilation system remained inactive. When occupancy was confirmed, input data were generated for each of the five ventilation modes (modes 0-4). Mode 0 represents the system-off state, while Modes 1-4 correspond with increasing levels of ventilation. The actual airflow rates were 8 m³/h for Mode 1, 30 m³/h for Mode 2, 60 m³/h for Mode 3, and 90 m³/h for Mode 4, which is the maximum output of the installed ventilation system. These values were determined based on the technical specifications and prior experimental calibration of the system. The ventilation levels do not increase in linear increments (either linear or 25% steps), but rather reflect practical operational thresholds. To ensure compatibility with the prediction model, all values were normalized to a 0-1 range using min–max scaling .

Step 3 involved predicting the CO_2 concentration for each airflow mode. The normalized input data was input into the CO_2 prediction model to estimate the indoor CO_2 concentration at time (t) for each

selected ventilation mode (i). The predicted values are denoted as $CO_{2\ pred\ i(t)}$. The model outputs a set of CO_2 predictions for all five ventilation modes, including the off state, as expressed in Eq. (1)

$$Outputdata = \left[CO_{2pred0(t)}, CO_{2pred1(t)}, CO_{2pred2(t)}, CO_{2pred3(t)}, CO_{2pred4(t)} \right] \quad (1)$$

In Step 4, the prediction errors were calculated by comparing each predicted value with the CO_2 concentration limit of 1,000 ppm ($CO_{2\ limit}$), as recommended by the WHO to ensure occupant comfort. The error for each airflow mode was computed using Eq. (2) as follows:

$$Error_i = CO_{2\ limit} - CO_{2\ pred\ i(t)} \quad (2)$$

where Error represents the error value, $CO_{2\ pred\ i}(t)$ the predicted indoor CO_2 concentration [ppm] at the next control time step (t+1), and $CO_2\ limit$ the predefined threshold set to 1,000 ppm. The subscript i represents the ventilation mode, where $i = [0,1,2,3,4]$, which corresponds to the off state and four increasing ventilation levels, respectively.

In Step 5, the control signal was selected and transmitted based on the computed errors. If none of the error values were negative, all the predicted CO_2 levels were considered to be below the threshold, while the ventilation system was turned off ($i=0$). If all the error values were negative, all predicted CO_2 concentrations were deemed to have exceeded the threshold, and the system switched to the

maximum airflow setting ($\dot{m}_{max}$, i = 4) for rapid air quality improvement. If one or more (but not all) error values were negative, the system selected the ventilation mode with the error closest to zero to minimize the energy consumption while maintaining air quality. The final control signal was then transmitted to the ventilation system, while the selected control settings were saved to a database.

Fig. 4 illustrates the algorithm-rule, which serves as the baseline control strategy in this study. This approach activates the ventilation system at predefined airflow levels whenever the current indoor CO_2 concentration exceeds a specified threshold. The threshold was set below 1,000 ppm, in accordance with the *Basic Classification of Indoor Air Quality* outlined in ASHRAE Standard 62.1 [36]. According to the standard, the difference between the indoor and outdoor CO_2 concentrations indicates the IAQ level: a difference of 400 ppm or less indicates high IAQ, 400-600 ppm medium IAQ, 600-1,000 ppm moderate IAQ, and above 1,000 ppm low IAQ. Based on these classifications, specific airflow levels were assigned to each CO_2 concentration range.

In this study, the mock-up and living lab environments utilized ventilation systems capable of delivering four levels of airflow (Modes 1-4), although the actual airflow rates varied based on the configuration of the system. The living lab, located at 'C' University in Dongjak-gu, Seoul, South Korea, covers an area of 62.83 m^2 with a volume of 169.64 m^3. The thermal transmittance (U-value) is 3.322 W/m^2·K for the roof and 0.463 W/m^2·K for the walls and floor, while the infiltration rate is 180 m^3/h. The facility is occupied by 9 individuals from 10:00 to 20:00 on weekdays, with no occupancy on weekends or holidays. The ventilation system used is an energy recovery ventilation system. This environment enables data collection on indoor and outdoor conditions, ventilation control, and occupant information. Detailed specifications for the airflow settings of each system are provided in Sections 3.1 and 3.2. As illustrated in Fig. 4, the algorithm-rule logic activates Mode 2 when CO_2 concentrations are between 1,000 and 1,200 ppm, Mode 3 when they are between 1,200 and 1,400 ppm, and Mode 4 when they are above 1,400 ppm. If the CO_2 concentration falls below 1,000 ppm, the system automatically turns off. Preliminary experiments indicated that the airflow in Mode 1 was insufficient to considerably reduce the indoor CO_2 concentration. Consequently, Mode 1 was excluded from the implemented control logic in the mockup and living lab settings.

3. Experimental configuration

3.1. Mock-up

To evaluate the performance of the data-driven indoor CO_2 concentration prediction model and integrated optimal ventilation control, a mock-up environment was developed for experimental validation. Table 1 presents an overview and configuration of the mockup, while Fig. 5 shows the photographs of the experimental test setup. The mock-up was installed on the second floor of University "C," located in Heukseok-dong, Dongjak-gu, Seoul, South Korea. The internal space measured 2.6 m (D) × 1.35 m (W) × 2.4 m (H). To facilitate occupant monitoring, the enclosure was constructed using transparent antistatic panels, allowing visibility from all sides except the floor. Although the mock-up was situated indoors, the windows and doors were kept open to minimize environmental differences between the test space and its surroundings, as enclosed buffer spaces (server and control zones) could affect internal CO_2 levels.

As presented in Table 1, the mockup enabled the real-time acquisition of environmental variables, occupant data, and ventilation system operation status. The setup included different sensors and data acquisition devices such as IoT-based environmental sensors (AW-1008-K-P), image-based sensors, a Raspberry Pi 4, an Arduino Uno microcontroller unit (MCU), and a dedicated data server. Three AW-1008-K-P sensors were positioned at a height of 1.2 m on the interior wall, a height used in previous studies to monitor CO_2 concentrations [37–39]. This height was selected as it approximates the average breathing zone of seated and standing occupants, corresponding with ASHRAE Standard 55-2017 [40], which specifies representative measurement heights at 0.6, 1.1, and 1.7 m for environmental variables. Preliminary measurements at 0.6, 1.2, and 1.7 m confirmed that 1.2 m adequately represents the average indoor CO_2 concentration in this environment. One image-based sensor was installed inside the mockup and another outside to detect occupant presence and activity. An image of the occupant activity detection experiment is shown in Fig. 5(b). All sensor data were transmitted to the database via the Arduino Uno MCU.

The ventilation system used in the mockup, as shown in Fig. 5(a), was custom-designed to fit the internal space illustrated in Fig. 6. To implement the control modes (off, outdoor air intake, and internal circulation), four dampers ((1)-1, (2)-2, (3)-3, (4)-4) and two fans for supply and exhaust ((2) supply fan and (3) exhaust fan) were installed, each controlled individually. The supply and exhaust fans (BFM17051B) installed at positions (2) and (3) were 24V DC fans measuring 170 mm (D) × 170 mm (W) × 51 mm (H), capable of generating a maximum airflow of 377.4 M^3/H at a static pressure of 13.54 mmH$_2$O. A high-performance filter was placed in front of the supply fan to minimize outdoor fine dust intake.

The damper operation based on the system control mode was as follows: in off mode, all dampers ((1)-1 and (1)-4) were closed. In outdoor air intake mode, dampers (1)-1 and (1)-2 were open, while dampers (1)-3 and (1)-4 remained closed. In internal circulation mode,
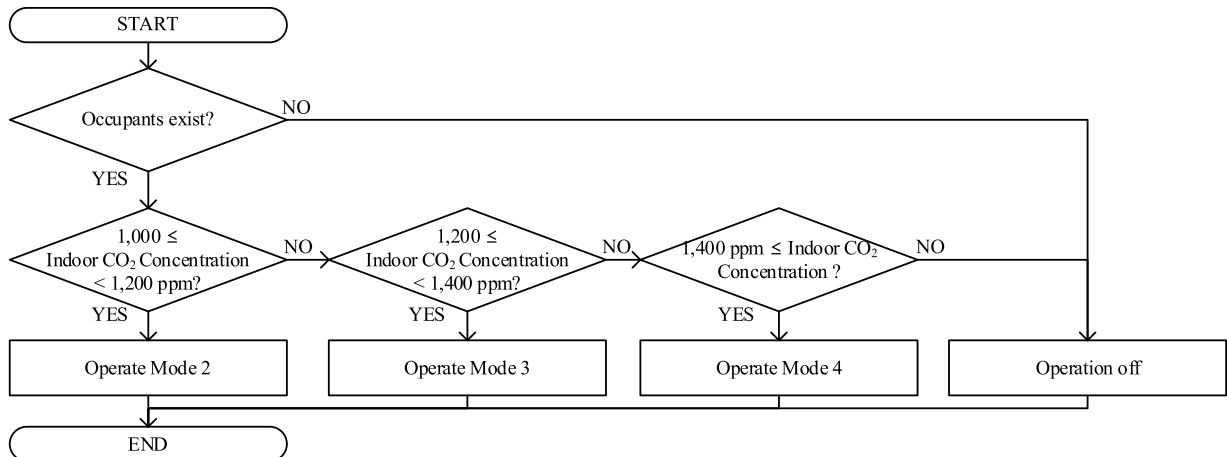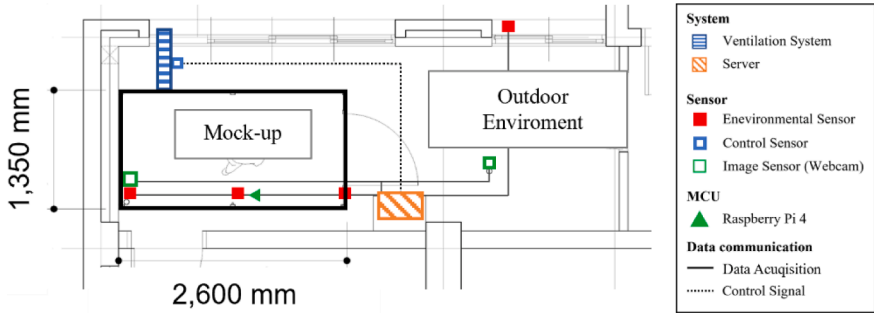

Fig. 4. Rule-based ventilation control algorithm.

Table 1
Mock-up overview.

Information	Details
Plan	
Location	'C' University, Heukseok-dong, Dongjak-gu, Seoul, South Korea
Area and volume	3.51 m^2 and 8.42 m^3, respectively
Envelope	Floor –Medium density fiber board
	Wall, roof –Uninterrupted acryl
Occupants	1 person
Ventilation system	Variable air volume system
	Mode 1: 8 M^3/H and Mode 2: 30 M^3/H
	Mode 3: 60 M^3/H and Mode 4: 90 M^3/H
Acquisition data	- CO$_2$ concentration [ppm]
	- Ventilation flowrate [M^3/H]
	- Number of occupants [person]
	- Occupant metabolic rate [MET]
	- BMI [kg/m^2]
	- Gender [0,1,2]

(a)

(b)
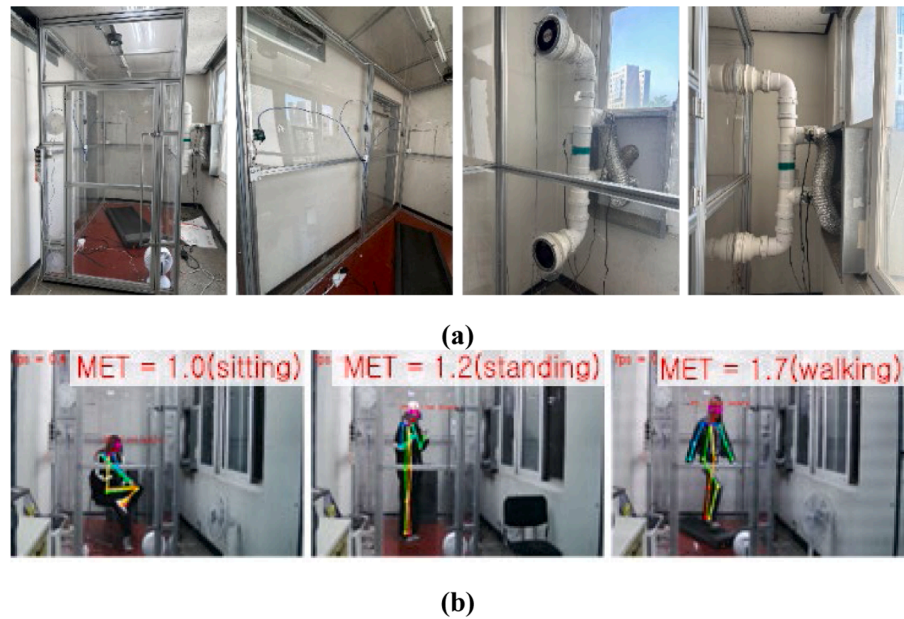
Fig. 5. (a) Mock-up test setup and (b) MET estimation.

dampers (1)-1 and (1)-2 were closed, while dampers (1)-3 and (1)-4 were open.

The experiments conducted with actual participants in the mockup comprised two main parts: 1) training data collection for the prediction model and 2) performance evaluation of the control algorithm. These experiments were conducted in February and March 2024. Participant details are presented in Table 2. Ten adult male and female participants, with BMIs ranging from 18.5-25 kg/m^2 were selected. They performed three activities in the mockup: sitting (1.0 MET), standing (1.2 MET), and walking (1.7 MET). Of the 10 participants, 8 (Subjects A–H) participated in the training data acquisition phase for the prediction model.

The CO$_2$ control performance of the ventilation algorithm was tested based on activity levels using 2 participants of different genders: Subject I (male) and Subject J (female). During the experiment, the activity levels were categorized as sitting (1.0 MET), standing (1.2 MET), and walking (1.7 MET). Both algorithm-rule and algorithm-opt were applied to evaluate their effectiveness. The experiment was conducted in two phases: the first 40 min utilized the algorithm-rule, followed by 40 min of optimal ventilation control, with each activity case lasting 80 min. A ventilation period of approximately 10 min was applied between each case.
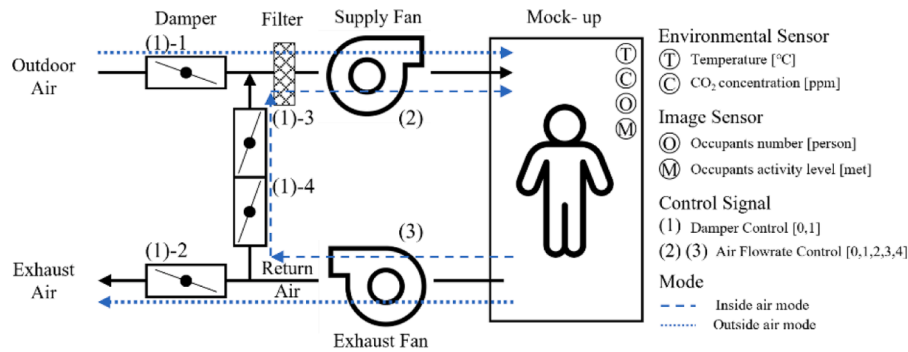
Fig. 6. Mock-up ventilation system.

Table 2
Participant details.

ID	Gender	Age (years)	BMI (kg/m^2)	Participation Period	Purpose of experiment
A	Male	32	22.89	02/12	Training data collection for prediction model
B	Male	28	18.81	02/13, 15	
C	Male	26	24.06	02/01, 02	
D	Male	24	22.6	02/08	
E	Female	28	18.26	02/14, 16	
F	Female	25	18.1	02/18, 19	
G	Female	28	19.18	02/03, 12	
H	Female	28	19.04	02/06, 07	
I	Male	33	22.31	03/13, 18, 19	Control algorithm evaluation
J	Female	25	20.06	3/17, 20, 24	

As shown in Fig. 5(b), MET values were calculated using images captured by video sensors, with joint data extracted via OpenPose. The key joints were identified in the images, and based on their positions, the MET values for the different activities (sitting, standing, and walking) were computed. OpenPose provided the joint positions, which were then processed to calculate the corresponding MET values for each activity.

3.2. Living lab

To verify whether the developed optimal ventilation control can maintain IAQ in different environments, a simulation evaluation was conducted in a larger, real-world office facility (living lab) beyond the mock-up setup. The simulation used baseline data acquired from the actual conditions of the living lab, thereby enabling the evaluation of indoor CO_2 levels and ventilation energy consumption under different control methods.

An overview of the living lab is presented in Table 3. Located on the first floor of University "C," the office accommodates 9 occupants. The living lab is equipped with an ERV system and an electric heat pump (EHP) for heating and cooling. The ventilation system can be adjusted from a low airflow of 200 M^3/H to a high of 1,000 M^3/H, with two units positioned near the windows for outdoor air intake. Three environmental sensors were installed at a height of 1.2 m to collect indoor data, while a pyranometer was installed outdoors to monitor the indoor and outdoor CO_2 concentrations, which were recorded in the database at 15-s intervals. Occupant count and MET data were collected using image sensors at 20-30 s intervals.

Simulations were conducted using Design Builder 7.0.2.003 and EnergyPlus V9.4.0. The building model was developed in Design Builder, while an HVAC system was configured using EnergyPlus. The occupant schedule reflected the actual living lab timetable. The outdoor CO_2 concentration was set at 421.6 ppm, representing the global average for January 2024, according to CO_2 Earth. The simulation spanned one week and accounted for daily occupancy variations,

including classes, meetings, extended hours, and remote work, which corresponds to the actual living lab conditions.

MET values for the simulation were calculated using a similar method applied in the mock-up experiments. OpenPose analyzed images captured by video sensors to determine joint positions, which were then used to classify activity levels: sitting (1.0 MET), standing (1.2 MET), and walking (1.7 MET). For each activity, the most frequent MET value over a 5-min interval was selected. During office hours, sitting (1.0 MET) was the predominant activity, and this was reflected in the simulation to represent the typical occupant activity.

Fig. 7. In a living lab environment, data were collected to validate the reliability and accuracy of the EnergyPlus simulation model against actual experiments, as shown in Fig. 8. From March 11-17, indoor and outdoor CO_2 concentration data were collected (Fig. 8(a)) and used to calibrate the simulation model [42,43]. During calibration, the ventilation system operated from 10:00 to 20:00 on weekdays, while the calibration performance was evaluated using three metrics: mean absolute percentage error (MAPE), coefficient of variation of root mean square error (CvRMSE), and coefficient of determination (R^2) [44,45].

Throughout the experimental period, the number of occupants varied daily owing to external activities, vacations, and remote work. This occupancy information was incorporated into the simulation model. A comparison between the simulation results and actual data is presented in the indoor CO_2 concentration plot in Fig. 8(a). The simulation demonstrated excellent accuracy in predicting indoor CO_2 concentrations, with a MAPE, CvRMSE, and R^2 of 6.01%, 7.53%, and 0.9717, respectively.

From March 18-24, weekday occupancy schedule data were collected over one week to use in the experiment. As shown in Fig. 8(b), occupants were considered between 9:00 and 19:00, with the average number of occupants per time period being approximately 5.4 on Days 1 to 4. Day 3 recorded the highest number of occupants at 6.9, while Day 5 had the lowest at 4.3. Simulation-based optimal ventilation system control and performance evaluations were performed using the acquired data.

4. Results analysis and discussion

4.1. Performance evaluation of indoor CO2 prediction model

Test data excluded from the training process were employed to evaluate the prediction performance. Table 4 presents the optimal hyperparameters and corresponding performance metrics for each model. While all three models share a similar number of input (eight) and output (one) neurons, they differ in the configuration of hidden neurons and hyperparameter settings. The evaluation was conducted using four metrics: mean absolute error (MAE), MAPE, CvRMSE, and R^2.

The DNN model comprised three hidden layers with 512, 128, and 176 neurons. A learning rate of 0.001 was applied, while ReLU was used as the activation function. This model achieved an MAE, CvRMSE, and
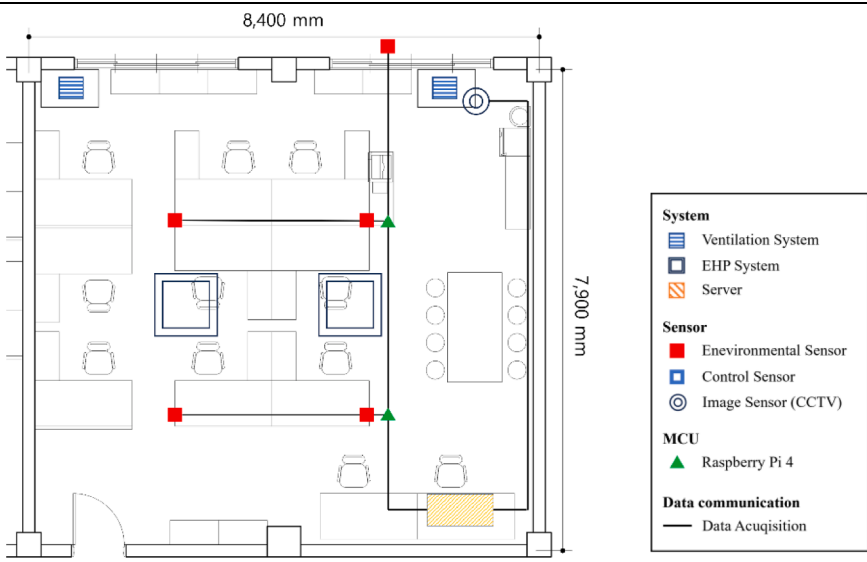
Table 3

Living lab overview.

Information	Details
Plan	
Location	'C' University, Dongjak-gu, Seoul, South Koreas
Area and volume	62.83 m^2 and 169.64 m^3, respectively
Thermal transmittance (U-value)	Roof: 3.322 W/m^2·k
	Wall, Floor: 0.463 W/ m^2·k
Infiltration	180 M^3/H
Occupants schedule	9 persons, 10:00-20:00, no occupants on weekends and holidays
Ventilation system	Energy recovery ventilation
	Mode 1: 200 M^3/H and Mode 2: 300 M^3/H
	Mode 3: 500 M^3/H and Mode 4: 1,000 M^3/H
Acquisition data	- CO_2 concentration [ppm]
	- Ventilation flowrate [M^3/H]
	- Number of occupants [person]
	- Occupant metabolic rate [MET]
	- BMI [kg/m^2]
	- Gender [0,1,2]

Fig. 7. Living-lab setup.

R^2 of 46.22 ppm, 14.40%, and 0.8242, respectively.

In the GRU model, the performance varied with the window size, with the best result obtained using a window size of 2 and two hidden layers comprising 384 and 512 neurons. The model employed a learning rate of 0.001 and Tanh activation function. This configuration yielded an MAE, CvRMSE, and R^2 of 32.97 ppm, 21.49%, and 0.6083, respectively.

The LSTM model demonstrated sensitivity to the window size, with an optimal performance at a window size of 2. This model included two hidden layers with 32 and 352 neurons, while Tanh was used as the activation function with a learning rate of 0.001. The LSTM model achieved the highest accuracy with an MAE, CvRMSE, and R^2 of 6.08 ppm, 4.75%, and 0.9809, respectively.

Overall, the LSTM outperformed the DNN and GRU models across all evaluation metrics owing to its capacity to effectively capture long-term dependencies inherent in time-series data. Hence, the LSTM model with a window size of 2 was selected as the prediction component for the proposed algorithm-opt.

As a preliminary study, this study focused on model simplicity and interpretability. To establish a clear baseline, a comparison was made using fundamental model architectures. Future studies will include more advanced architectures—attention-based models and CNN-LSTM
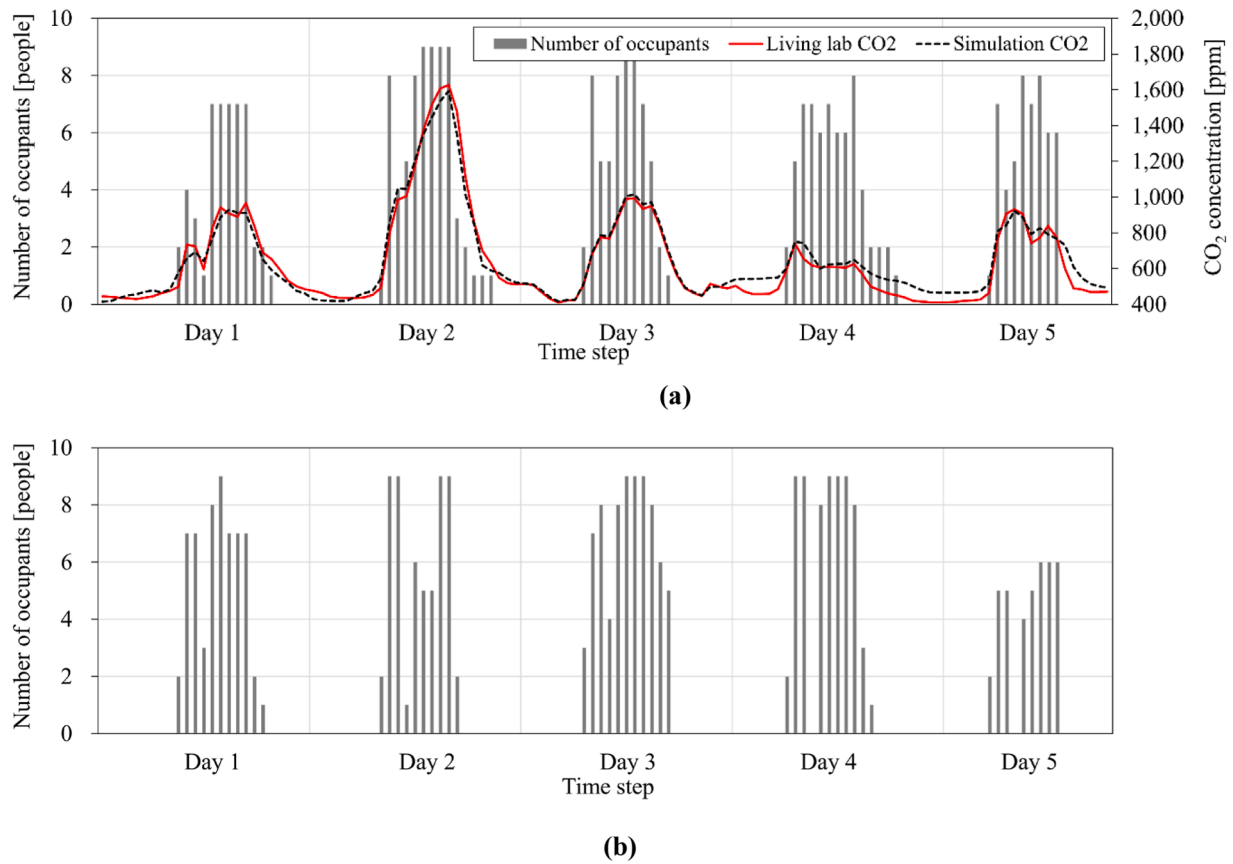
(a)

(b)

Fig. 8. Data acquisition in the living lab: (a) data for the simulation model calibration and (b) occupant schedule during the experiment period.

Table 4
Prediction accuracy.

Model	Window size	Structure of hidden layers	Learning rate	Activation function	MAE (ppm)	CvRMSE (%)	R²
DNN	-	512-128-176	0.001	ReLU	46.22	14.40	0.8242
GRU	2	384-512	0.001	Tanh	32.97	21.49	0.6083
	3	320-512	0.001	Tanh	33.45	21.58	0.6051
	4	256-512	0.001	Tanh	38.03	22.27	0.5793
LSTM	2	32-352	0.001	Tanh	6.08	4.75	0.9809
	3	32-384	0.001	Tanh	17.78	8.61	0.9372
	4	32-224	0.001	Tanh	25.70	12.72	0.8627

networks—to further enhance prediction performance. GRU was included in the analysis owing to its faster training speed and lower computational costs compared to LSTM, making it more suitable for environments with Its simpler architecture also translates to reduced training time and computational cost the potential application in real-time ventilation control systems, GRU was evaluated to assess its feasibility and effectiveness in practical deployment scenarios [46–48].

4.2. Comparison of indoor CO2 comfort ratio

To assess the effectiveness of the optimal ventilation control strategy, a two-phase experiment was conducted in the mockup and living lab settings. The evaluation criterion was the proportion of time during which IAQ remained within the comfort range, defined as the period when the indoor CO_2 concentration did not exceed 1,000 ppm. This percentage was calculated as the ratio of the duration during which CO_2 levels were maintained below 1,000 ppm to the total duration of the control period.

In the mock-up experiment, the effect of each control algorithm on the indoor CO_2 concentration was analyzed based on the occupant activity levels, as shown in Fig. 9. For each activity, the proportion of time during which the CO_2 concentration remained below the threshold of 1,000 ppm was quantified as the indoor air comfort ratio, as illustrated in Fig. 10. Fig. 9(a), (c), and (e) show the results of applying algorithm-rule under sitting, standing, and walking conditions, respectively, while Fig. 9(b), (d), and (f) show those for similar activities using the proposed algorithm-opt. Each plot shows the experimental results for Subjects I (male) and J (female). The data represents the actual CO_2 concentration measured in the controlled environment, including the control mode of the ventilation system. For algorithm-opt, the predicted CO_2 concentration, denoted as $CO_{2\ pred\ i(t)}$, is also presented, where i represents the mode number and t the time in minutes. The predictions are made 5 min ahead. The predicted CO_2 values shown in the plots indicate the estimated concentration 5 min ahead, which corresponds with the optimal airflow mode determined by the algorithm described in Section 2.2.

The results for sitting (1.0 MET) are shown in Fig. 9(a) and 9(b). Under algorithm-rule, indoor CO_2 concentrations exceeded 1,000 ppm, triggering control activation at the next scheduled intervals (at 15 min for Subject I and 20 for Subject J), both at Mode 2. As regards Subject I,
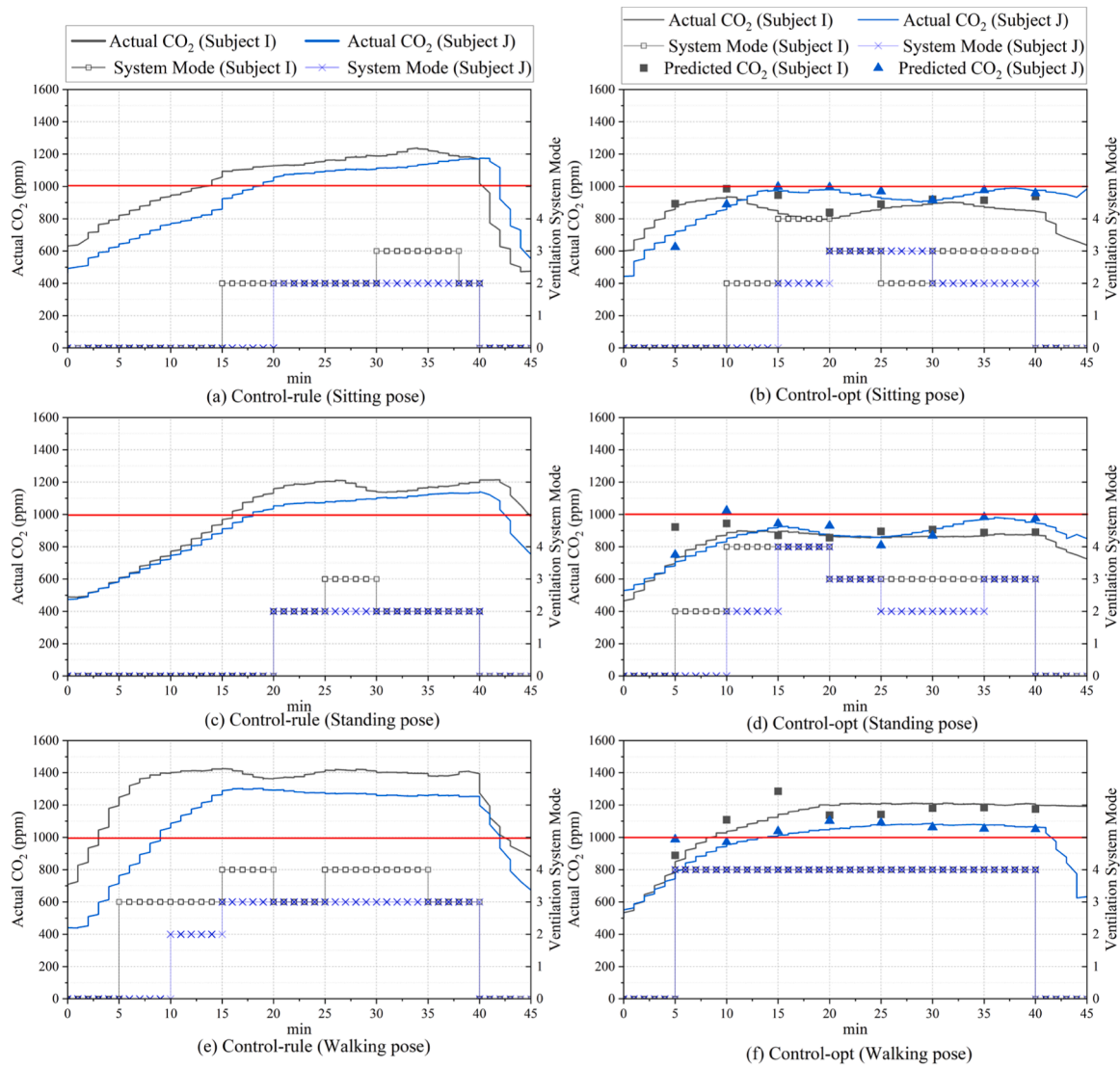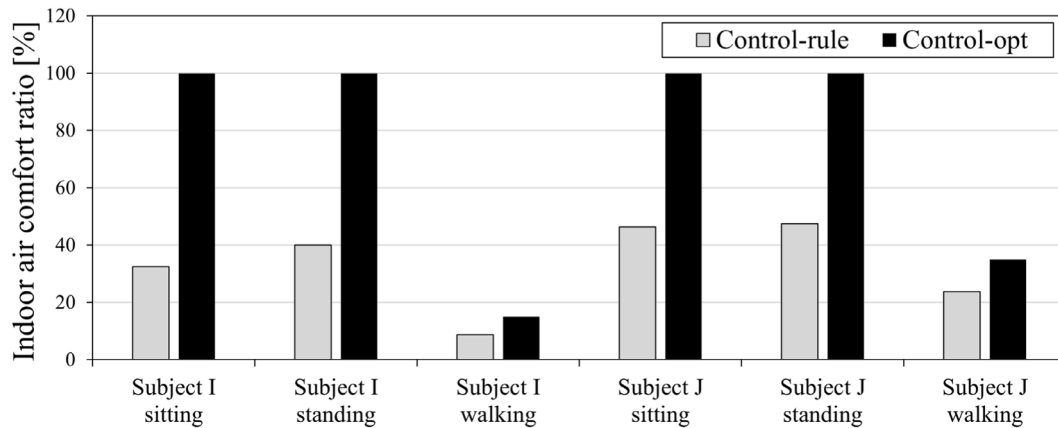
Fig. 9. Mock-up experiment results.

Fig. 10. Indoor comfort ratio analysis based on mock-up experiments.

CO_2 concentration exceeded 1,200 ppm at 30 min, prompting a switch to Mode 3. As shown in Fig. 10, the indoor air comfort ratios for Subjects I and J were 32.5% and 46.25%, respectively. Conversely, algorithm-opt continuously predicted CO_2 levels to proactively adjust the airflow and maintain concentrations below 1,000 ppm. At the start of the experiment, both subjects exhibited low predicted CO_2 values, while the ventilation system was turned off. For Subject I, Mode 2 was activated at 10 min when the predicted CO_2 concentration was $CO_{2\ pred\ 2(t=5)} = 985$

ppm, while for Subject J, control started at 15 min with $CO_{2\ pred\ 2(t=5)}$ = 998 ppm. Throughout the experiment, CO_2 concentrations for both subjects remained below 1,000 ppm, thereby resulting in indoor air comfort ratios of 100% (Fig. 10).

As regards standing (1.2 MET), using the algorithm-rule (Fig. 9(c)) resulted in both subjects exhibiting CO_2 concentrations exceeding 1,000 ppm at 15 min, with ventilation operated in Modes 2 and 3 thereafter. However, the average exceedance period was 23 min. The indoor air comfort ratios for Subjects I and J were 40% and 47.5%, respectively.

Under algorithm-opt (Fig. 9(d)), Subject I triggered Mode 2 operation at 5 min with a predicted value of $CO_{2\ pred\ 2(t=5)}$ = 922 ppm, followed by Mode 4 at 10 min with $CO_{2\ pred\ 4(t=5)}$ = 944 ppm. The average CO_2 concentration was 862 ppm, marking the lowest across all the experimental conditions. Subject J activated Mode 2 at 10 min and Mode 4 at 15 min, alternating between Modes 2 and 3 to maintain CO_2 concentrations below 1,000 ppm. No threshold exceedance occurred during the control period for either subjects, thereby resulting in an indoor air comfort ratio of 100%.
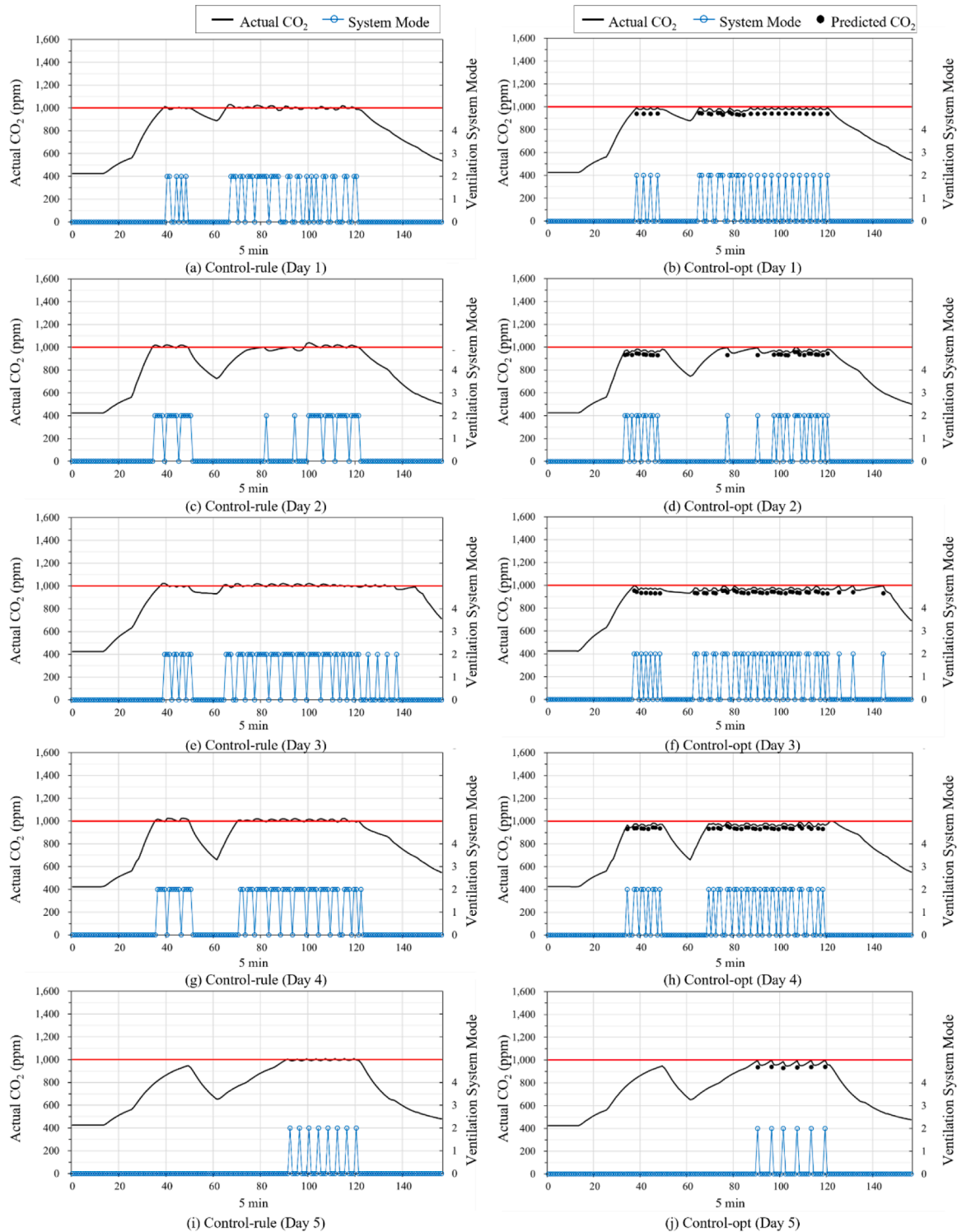
Fig. 11. Ventilation control performance in the living lab.

Regarding walking (1.7 MET), algorithm-rule (Fig. 9(e)) resulted in CO_2 concentrations above 1,000 ppm within the first 10 min for both subjects. Despite subsequent operation in Mode 3 or higher, the concentrations remained between 1,100 and 1,200 ppm, peaking above 1,400 ppm. The average CO_2 levels were 1,327 and 1,258 ppm for Subjects I and J, respectively, thereby resulting in indoor air comfort ratios of 8.75% and 23.75%, respectively. Under optimal control (Fig. 9 (f)), Subjects I and J had a predicted value of $CO_{2\ pred\ 4(t=5)} = 890$ ppm and $CO_{2\ pred\ 4(t=5)} = 982$ ppm at 5 min, respectively, thereby triggering the immediate operation of Mode 4, which was maintained until the end of the experiment. The resulting indoor air comfort ratios improved to 13.95 and 36.65% for Subjects I and J, respectively (Fig. 10). Although CO_2 levels exceeded 1,000 ppm for the majority of the time, the average concentrations under optimal control were approximately 1,005 ppm compared to 1,365 ppm under algorithm-rule—representing a 26% decrease and demonstrating enhanced IAQ.

In summary, the proposed optimal ventilation control system outperformed the conventional rule-based control method in maintaining indoor CO_2 comfort, particularly under low MET conditions such as sitting (1.0 MET) and standing (1.2 MET), wherein considerable improvements in the indoor air comfort ratio were observed. However, maintaining CO_2 levels below 1,000 ppm during walking (1.7 MET) proved more challenging owing to the limited airflow capacity of the ventilation system relative to the increased CO_2 generation from higher physical activity. To address this, improvements such as integrating supplemental air-handling systems and advancing adaptive control logic capable of relearning occupant activity patterns and IAQ dynamics are required.

Across all the experiments, Subject J (female) consistently emitted less CO_2 than Subject I (male), thereby suggesting a gender-based difference in CO_2 generation likely attributable to physiological factors (differences in respiratory volume and metabolic rate) [41]. Generally, males have higher metabolic rates and larger body sizes, which contribute to increased respiratory CO_2 output. Under algorithm-rule (Fig. 9), Subject I maintained higher CO_2 levels than Subject J. Under optimal control, the CO_2 concentrations were regulated to similar levels for the subjects across all activity levels, except for walking. This suggests that the developed method effectively predicts CO_2 levels and adjusts ventilation regardless of gender, thereby maintaining indoor CO_2 concentrations within the comfort range for different occupants.

The second performance evaluation was conducted using simulations in a living lab environment, wherein algorithm-opt and -rule control strategies were applied from 10:00-20:00. The analysis examined CO_2 concentration variations and energy consumption patterns. The results over 5 d are shown in Fig. 11, with the left plot showing the results under algorithm-rule and right those under algorithm-opt.

Under algorithm-rule, CO_2 concentrations rapidly increased in the morning as occupant numbers increased, exceeding 1,000 ppm before the system activated. This resulted in extended periods of CO_2 levels, with the reactive nature of the system leading to slower ventilation responses after the threshold was exceeded. Conversely, the optimal ventilation control system proactively activated the ventilation system 5 min ahead whenever the predicted CO_2 concentration was projected to exceed 1,000 ppm, thereby preventing the concentration from surpassing the threshold.

The system activation and occupancy patterns were analyzed daily. On Day 5 (Fig. 11(i) and 11(j)), the CO_2 concentrations were maintained below 1,000 ppm, with notably fewer system activations. This can be attributed to the reduced number of occupants compared to those on other days (Fig. 8(b)), which slowed the CO_2 accumulation rate. On Day 3 (Fig. 11(e) and 11(f)), when occupants remained until 20:00, CO_2 levels remained high until later in the day, thereby resulting in an extended system operation.

Fig. 12 shows the proportion of time during which the CO_2 concentrations remained below 1,000 ppm across all occupancy periods. For algorithm-rule, the time exceeding the CO_2 threshold increased with an increase in the number of occupants. On Day 3, with the highest average occupancy of 6.9 people, the indoor comfort ratio was 64%, while on Day 5, with the lowest occupancy, it reached 95%. These variations highlight the differences in performance even under a similar control method. However, algorithm-opt successfully maintained CO_2 concentrations below 1,000 ppm throughout the experimental period, demonstrating its ability to predict and respond to CO_2 fluctuations driven by occupancy changes. This confirms that algorithm-opt, which incorporates occupancy variations, provides superior air quality control than the conventional method.

4.3. Energy consumption analysis

This section presents the investigation of the impact of the optimal ventilation control on the ventilation energy consumption. Energy consumption was calculated based on the power consumption of the ventilation system fan, with the fan operating at 24 V and 5 A, thereby resulting in a maximum output power of 14.4 W. To calculate the energy consumption at different airflow rates, the performance load factor (PLF) from EnergyPlus was applied, as expressed in Eq. (3). The constant, linear, and higher-order coefficients (2nd–4th) were derived from the inlet vane damper coefficients provided by EnergyPlus.

$$PLF = C_1 + C_2 \cdot FF + C_3 \cdot FF^2 + C_4 \cdot FF^3 + C_5 \cdot FF^4 \qquad (3)$$

where PLF represents the energy consumption coefficient and FF the airflow ratio. The constants C_1, C_2, C_3, C_4, and C_5 represent empirical curve coefficients in a polynomial equation used to approximate the part load performance (PLF) of HVAC equipment as a function of the flow
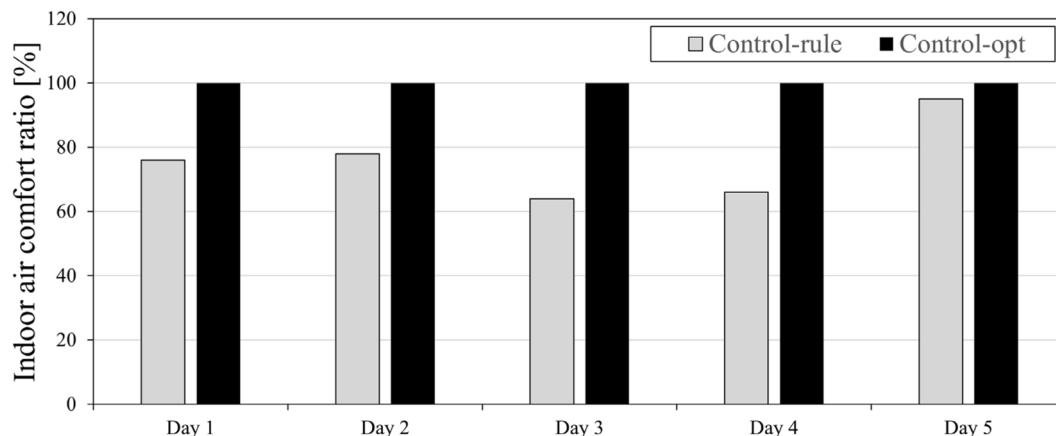

Fig. 12. Indoor comfort ratio analysis for the living lab.

fraction (FF). These coefficients are obtained from manufacturer data or EnergyPlus simulation parameters. For this experiment, the EnergyPlus settings were used to calculate the required values.

Using the calculation method, the energy consumptions of algorithm-rule and -opt were compared. Fig. 13 shows the energy consumption and ventilation system operation time for each control method in the mock-up and living lab environments. In the mock-up environment, the energy consumption of the ventilation system for each airflow mode was calculated as follows: Mode 1 at 11.17 W, Mode 2 at 13 W, Mode 3 at 17.85 W, and Mode 4 at 28.51 W. Based on the mock-up results shown in Fig. 13(a), Subject I consumed between 3.65 and 11.61 Wh in algorithm-rule, while the optimal ventilation control resulted in energy consumptions ranging from 7.52-13.38 Wh. Depending on the activity, algorithm-opt consumed more energy (between 15 and 182%). Subject J exhibited a similar trend, with algorithm-opt consuming more energy (between 43 and 90%) than algorithm-rule.

The greatest difference in energy consumption between the control methods was observed during the standing activity. For Subject I, the optimal ventilation control resulted in an approximately 75% longer system operation time than algorithm-rule (20-35 min), while Subject J exhibited a 33% increase (20-30 min). During sitting and walking, optimal ventilation control performed proactive control, thereby

predicting CO_2 levels one time-step ahead. However, during standing, the system overestimated the CO_2 increase rate, thereby resulting in excessive control two to three time-steps ahead. Consequently, the ventilation system operated more frequently than algorithm-rule. This overestimation is attributed to limited training data for standing, thereby preventing the model from accurately capturing the CO_2 changes during this activity. Compared to sitting and walking, standing appears to require extensive data to improve the CO_2 prediction accuracy. Hence, future studies must increase the training data for standing and enhance the model performance.

The ventilation system operation time analysis, represented by a line graph in Fig. 13(a), revealed that algorithm-opt resulted in longer operation times than algorithm-rule. For both subjects, the operation time increased with an increase in the activity levels, attributable to the optimal control ability of the system to adjust ventilation in response to elevated CO_2 concentration rates during intense activity. Further, the difference in the operation time between the control methods varied by gender. Even under similar activity conditions, Subject I (male) required longer system operation times than Subject J (female). This indicates that algorithm-opt, using the LSTM network, incorporates historical data patterns to predict CO_2 levels and adjust ventilation based on the gender-based differences in CO_2 emissions.

In a living lab environment, the energy consumption of the

	Sub. I sitting	Sub. I standing	Sub. I walking	Sub. J sitting	Sub. J stading	Sub. J walking
Algorithm-rule	5.81 Wh	3.65 Wh	11.61 Wh	4.32 Wh	4.32 Wh	10.04 Wh
Algorithm-opt	7.52 Wh	10.31 Wh	13.38 Wh	6.22 Wh	8.22 Wh	12.50 Wh

(a)

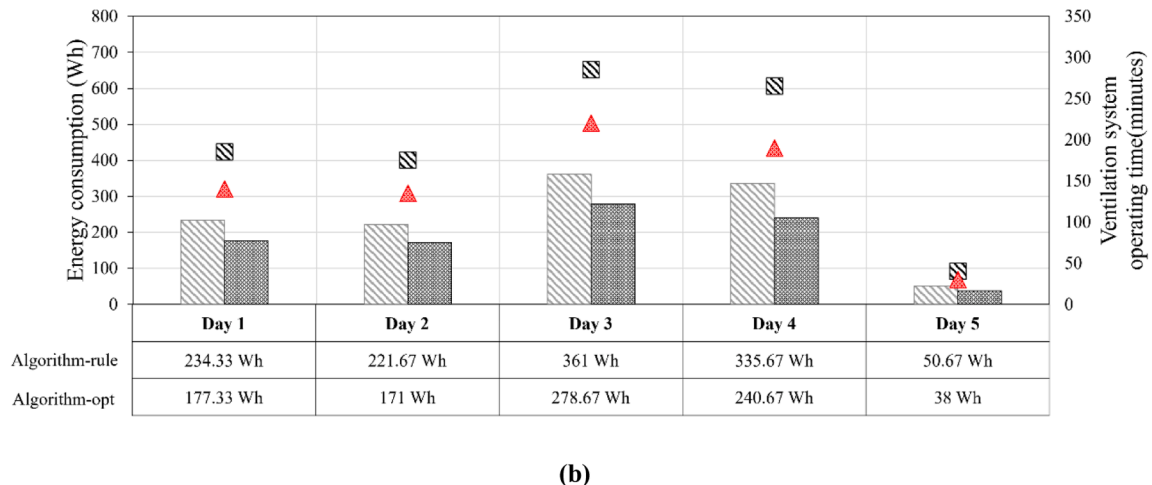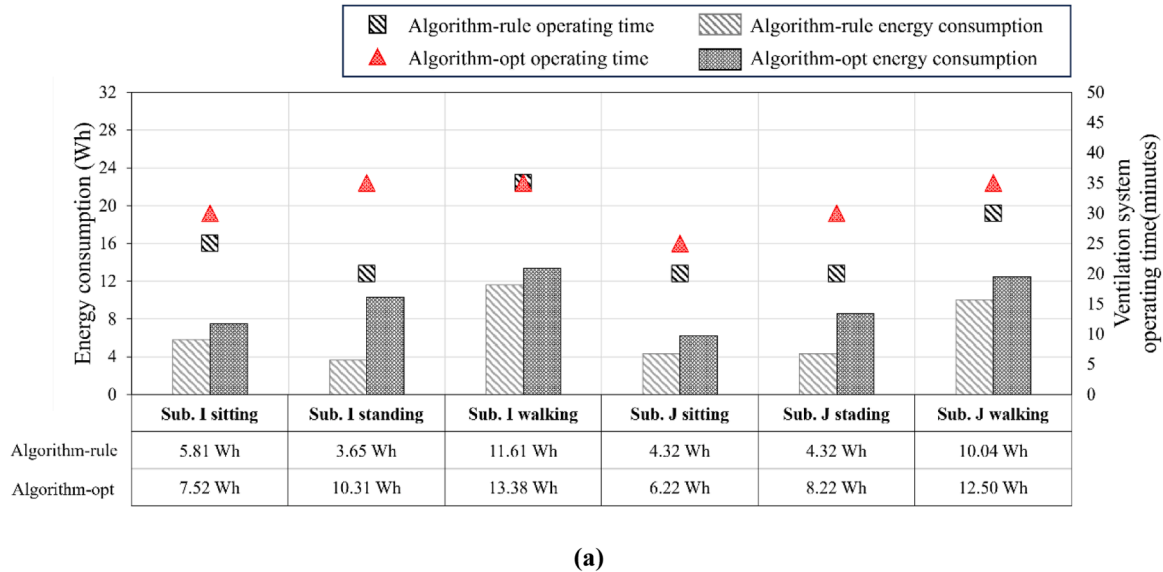|	Day 1	Day 2	Day 3	Day 4	Day 5
Algorithm-rule	234.33 Wh	221.67 Wh	361 Wh	335.67 Wh	50.67 Wh
Algorithm-opt	177.33 Wh	171 Wh	278.67 Wh	240.67 Wh	38 Wh

(b)

Fig. 13. Energy consumption results: (a) Mock-up and (b) living lab.

ventilation system fan was measured under similar conditions to those in the mock-up setting. The energy consumption for each airflow mode was calculated as 58, 76, 137, and 395 W for Modes 1-4, respectively, as shown in Fig. 13(b). Throughout the experiment, both control methods operated at Mode 2 airflow, as shown in Fig. 11. Over the 5-d testing period, the total energy consumption was 1,203.33 Wh for algorithm-rule and 905.67 Wh for optimal ventilation control, indicating energy savings of approximately 24.74% using the proposed method.

Considerable differences were observed on Days 3 and 5. On Day 3, owing to high occupancy density, algorithm-rule recorded the highest energy at 361 Wh, while optimal control reduced this to 278.67 Wh by accurately predicting the CO_2 fluctuations. On Day 5, with lower occupancy and stable outdoor CO_2 concentrations, the optimal control further reduced consumption to 240.67 Wh, compared to 335.67 Wh under algorithm-rule. These differences were primarily owing to the shorter system operation durations under optimal control conditions. From Days 1-5, the operation time for algorithm-rule ranged 40-185 min, while it ranged 30-140 min for optimal control. This reduction in the operating time is attributed to the effectiveness of algorithm-opt in predicting CO_2 concentration changes and minimizing unnecessary ventilation.

Consequently, the optimal ventilation control strategy effectively reduced energy consumption while maintaining stable IAQ. The differing energy consumption trends between the mock-up experiment and living lab simulation are attributed to variations in the ventilation system capacity and indoor spatial scale. The mock-up, designed for a single occupant, had a limited ventilation capacity and more confined space, thereby making CO_2 levels more volatile and requiring frequent changes in the operation mode, which increased energy consumption. Conversely, the living lab featured a larger system capacity and spatial volume, thereby resulting in an average CO_2 fluctuation per occupant and reduced operational demands. These findings suggest that the proposed control method is more energy-efficient in larger environments with sufficient ventilation capacity. Hence, future implementations should be optimized based on the space scale and system performance.

4.4. Discussion

This study developed a real-time CO_2 concentration prediction model that incorporates occupant information and applied an optimized ventilation control system based on the predicted values. While previous studies have proposed ventilation control systems that consider occupant behavior and environmental variables, the majority rely on reactive methods that only activate ventilation after IAQ exceeds the threshold. Such methods are limited by delayed responses once the IAQ is breached [12,13]. This study proposes a predictive ventilation control system that anticipates IAQ deterioration and proactively adjusts the system in real-time, thereby enabling more effective control than conventional methods.

Previous studies focused on identifying the current occupancy state and activity level to control ventilation systems. Choi et al [16]. used video-based occupant sensors to monitor occupancy and regulate an ERV system, thereby achieving up to 40% energy savings. However, such studies relied on reactive control, thereby activating the ventilation system only after CO_2 concentrations exceeded predefined thresholds. This study utilizes real-time CO_2 concentration predictions to proactively adjust the ventilation system, thereby improving the IAQ and energy efficiency simultaneously. Kim et al [17]. developed a multi-deep Q-network algorithm that identified basic occupancy states (work, sleep, and absence) to control cooling, ventilation, and humidification systems, thereby achieving energy savings between 6.3 and 21%. While both approaches utilize occupant data to optimize ventilation control, this study advances this concept by integrating personalized occupant information (MET, BMI, and gender) into the real-time CO_2 concentration prediction model. This integration enables a predictive control strategy that not only anticipates air quality deterioration but also

balances IAQ and energy consumption. Further, by factoring in diverse physiological traits, the proposed method offers a more refined and responsive control system than previous studies

The analysis of the results further confirms the effectiveness of the proposed system in optimizing ventilation control. The optimized ventilation control algorithm (algorithm-opt) successfully reduced energy consumption by 24.74% while maintaining stable indoor CO_2 concentrations in a mock-up environment. This highlights the potential of dynamically adjusting ventilation operations to simultaneously optimize IAQ and energy consumption. Unlike the rule-based control method, which activates the ventilation system only after CO_2 levels exceed predefined thresholds, algorithm-opt predicts CO_2 concentrations and proactively adjusts the system in real-time, effectively saving energy and improving IAQ.

The multi-occupant living lab simulation demonstrated that algorithm-opt reduced the ventilation system operation time and managed energy consumption efficiently. Although algorithm-opt consumed more energy in the mock-up environment, it maintained stable CO_2 concentrations and optimized the ventilation system. In the living lab environment, accurate predictions reduced the system operation time, thereby resulting in a 24.74% decrease in energy consumption. These results highlight the practicality and efficiency of the prediction-based ventilation control strategy, which simultaneously addresses both IAQ and energy consumption.

Future research will focus on expanding and diversifying the training dataset to improve the accuracy of CO_2 concentration predictions based on occupant behavior and activity levels. Additionally, integrating adaptive learning techniques for real-time model retraining will ensure system responsiveness to changing indoor conditions, thereby enabling accurate IAQ management. While this study treated outdoor CO_2 concentrations as relatively insignificant, outdoor CO_2 may be crucial in diluting indoor CO_2 concentrations in other environments. Future research should develop models that consider changes in outdoor CO_2 concentrations to improve prediction accuracy and overall air quality control. Furthermore, advanced model architectures (Attention-based models, CNN-LSTM, ResNet + LSTM, and other hybrid models) will be examined to further improve prediction performance and effectively address the complexities of real-time ventilation control.

5. Conclusions

This study developed a personalized, real-time CO_2 concentration prediction model and implemented an optimized ventilation control system to improve IAQ. The proposed system offers a prediction-based ventilation control approach that anticipates and resolves issues associated with reactive control methods.

The LSTM model exhibited superior prediction performance than other models, achieving the lowest MAE of 6.08 ppm, CvRMSE) of 4.75%, and R^2 value of 0.9809. This demonstrates a considerably higher prediction accuracy than the DNN and GRU models.

The optimized ventilation control system (algorithm-opt) effectively maintained stable indoor CO_2 concentrations and enhanced IAQ. In the mock-up environment, algorithm-opt successfully kept CO_2 concentrations below the 1,000-ppm threshold, thereby improving IAQ. Although CO_2 concentrations occasionally exceeded 1,000 ppm during walking activities, algorithm-opt considerably reduced CO_2 levels than the rule-based control system, thereby demonstrating the effectiveness of the prediction-based control approach.

The optimized control system (algorithm-opt) demonstrated notable improvements in energy efficiency. In the mock-up environment, algorithm-opt achieved a 24.74% reduction in energy consumption while maintaining stable indoor CO_2 concentrations. However, the mock-up experiment revealed that algorithm-opt consumed more energy than the rule-based control system, which can be attributed to the increased fan usage required to improve IAQ.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this study, the authors used OpenAI's ChatGPT to improve the language and readability. Further, the authors carefully reviewed and edited the content to ensure accuracy and take full responsibility for the content of the manuscript.

Funding

CRediT authorship contribution statement

Kang Woo Bae: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Conceptualization. **Eun Ji Choi:** Validation, Resources, Data curation. **Young Jae Choi:** Resources, Methodology, Formal analysis, Data curation. **Ji Young Yun:** Investigation, Data curation. **Geun Young Yun:** Data curation. **Hyeun Jun Moon:** Methodology, Data curation. **Jin Woo Moon:** Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] A. Cincinelli, T. Martellini, Indoor air quality and health, Int. J. Environ. Res. Public Health 14 (11) (2017) 1286, https://doi.org/10.3390/ijerph14111286.

[2] K.E. Paleologos, M.Y. Selim, A.O. Mohamed, Indoor air quality: pollutants, health effects, and regulations, in: D. Tsoutsos (Ed.), Pollution Assessment for Sustainable Practices in Applied Sciences and Engineering, Butterworth-Heinemann, Oxford, 2021, pp. 405–489.

[3] P. Kumar, B. Imam, Footprints of air pollution and changing environment on the sustainability of built infrastructure, Sci. Total Environ. 444 (2013) 85–101, https://doi.org/10.1016/j.scitotenv.2012.11.056.

[4] A.J. Koivisto, K.I. Kling, O. Hänninen, M. Jayjock, J. Löndahl, A. Wierzbicka, A. S. Fonseca, K. Uhrbrand, B.E. Boor, A.S. Jiménez, Source specific exposure and risk assessment for indoor aerosols, Sci. Total Environ. 668 (2019) 13–24, https://doi.org/10.1016/j.scitotenv.2019.02.398.

[5] A. Persily, L. de Jonge, Carbon dioxide generation rates for building occupants, Indoor Air 27 (5) (2017) 868–879, https://doi.org/10.1111/ina.12383.

[6] S. García, M. Monserrat-Mesquida, S. Mas-Fontao, E. Cuadrado-Soto, M. Ortiz-Ramos, P. Matía-Martín, L. Daimiel, C. Vázquez, J.A. Tur, C. Bouzas, Body composition and CO2 dietary emissions, Front. Public Health 12 (2024) 1432109, https://doi.org/10.3389/fpubh.2024.1432109.

[7] W. Ji, M. Luo, B. Cao, Y. Zhu, Y. Geng, B. Lin, A new method to study human metabolic rate changes and thermal comfort in physical exercise by CO2 measurement in an airtight chamber, Energy Build 177 (2018) 402–412, https://doi.org/10.1016/j.enbuild.2018.08.018.

[8] C.A. Harms, S. Rosenkranz, Sex differences in pulmonary function during exercise, Med. Sci. Sports Exerc. 40 (4) (2008) 664–668, https://doi.org/10.1249/mss.0b013e3181621325.

[9] M. del Mar Castilla, J.D. Álvarez, J.E. Normey-Rico, F. Rodríguez, M. Berenguel, A multivariable nonlinear MPC control strategy for thermal comfort and indoor-air quality, in: Proc. IECON 2013 –39th Annu. Conf. IEEE Ind, Electron. Soc., 2013, pp. 7908–7913, https://doi.org/10.1109/IECON.2013.6700454.

[10] A. Berouine, R. Ouladsine, M. Bakhouya, M. Essaaidi, Towards a real-time predictive management approach of indoor air quality in energy-efficient buildings, Energies 13 (12) (2020) 3246, https://doi.org/10.3390/en13123246.

[11] W. Zhang, Y. Wu, J.K. Calautit, A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment, Renew. Sustain. Energy Rev. 167 (2022) 112704, https://doi.org/10.1016/j.rser.2022.112704.

[12] Z. Wang, J. Calautit, P.W. Tien, S. Wei, W. Zhang, Y. Wu, L. Xia, An occupant-centric control strategy for indoor thermal comfort, air quality and energy management, Energy Build 285 (2023) 112899, https://doi.org/10.1016/j.enbuild.2023.112899.

[13] C. Turley, M. Jacoby, G. Pavlak, G. Henze, Development and evaluation of occupancy-aware HVAC control for residential building energy efficiency and occupant comfort, Energies 13 (2020) 5396, https://doi.org/10.3390/en13205396.

[14] S. Hong, S. Yeon, B. Seo, B.H. Yu and K.H. Lee, '"Variations of PMV based thermal comfort and cooling/heating load according to MET," KIEAE journal, vol. 17, no. 6, pp. 39-44. 10.12813/kieae.2017.17.6.039.

[15] J. Wang, J. Huang, Z. Feng, S. Cao, F. Haghighat, Occupant-density-detection based energy efficient ventilation system: Prevention of infection transmission, Energy Build 240 (2021) 110883, https://doi.org/10.1016/j.enbuild.2021.110883.

[16] Y.J. Choi, E.J. Choi, J.Y. Byun, H.J. Moon, M.K. Sung, J.W. Moon, CO2- and PM2.5-Focused Optimal Ventilation Strategy Based on Predictive Control, Indoor Air. 2025 (2025) 6652442. https://doi.org/10.1155/ina/6652442.

[17] S.H. Kim, Y.R. Yoon, J.W. Kim, H.J. Moon, Reinforcement Learning for Integrated Optimal Control of Ventilation System and an Air Purifier to Maintain Healthy IAQ, J. Korean Soc. Living Environ. Syst. 29 (2022) 176, https://doi.org/10.3390/buildings13123062.

[18] I. Mutis, A. Ambekar, V. Joshi, Real-time space occupancy sensing and human motion analysis using deep learning for indoor air quality control, Autom. Constr. 116 (2020) 103237, https://doi.org/10.1016/j.autcon.2020.103237.

[19] J. Skön, M. Johansson, M. Raatikainen, K. Leiviskä, M. Kolehmainen, Modelling indoor air carbon dioxide (CO2) concentration using neural network, Methods 14 (2020) 15–16.

[20] J.C.P. Putra, S. Safrilah, M. Ihsan, The prediction of indoor air quality in office room using artificial neural network, IOP Conf. Ser. Mater. Sci. Eng. 1977 (2021) 012045, https://doi.org/10.1063/1.5042896.

[21] Y. Zhu, S.A. Al-Ahmed, M.Z. Shakir, J.I. Olszewska, LSTM-based IoT-enabled CO2 steady-state forecasting for indoor air quality monitoring, Electronics 12 (2023) 107, https://doi.org/10.3390/electronics12010107.

[22] S. Liu, X. Liu, Q. Lyu, F. Li, Comprehensive system based on a DNN and LSTM for predicting sinter composition, Appl. Soft Comput. 95 (2020) 106574, https://doi.org/10.1016/j.asoc.2020.106574.

[23] Y.Wang Liu, X. Yang, L. Zhang, Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models, IEEE Access 6 (2018) 36873–36882, https://doi.org/10.1109/ITSC.2017.8317886.

[24] S. Jo, J. Kim, S. Kim, J. Youn, A comparative study on the performance of air quality prediction model based on DNN and LSTM, Proc. Korean Inst. Electr. Eng. Conf. (2019) 577–579, https://doi.org/10.3745/PKIPS.y2020m05a.577.

[25] X. Wang, J. Yan, X. Wang, Y. Wang, Air quality forecasting using the GRU model based on multiple sensors nodes, IEEE Sens. Lett. 7 (2023) 1–4, https://doi.org/10.1109/LSENS.2023.3290144.

[26] Q. Liao, M. Zhu, L. Wu, X. Pan, X. Tang, Z. Wang, Deep learning for air quality forecasts: a review, Curr. Pollut. Rep. 6 (2020) 399–409, https://doi.org/10.1007/s40726-020-00159-z.

[27] F.M. Shiri, T. Perumal, N. Mustapha, R. Mohamed, A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU, arXiv preprint arXiv:2305.17473. 10.48550/arXiv.2305.17473.

[28] A. Klami, S. Virtanen, S. Kaski, Bayesian canonical correlation analysis, J. Mach. Learn. Res. 14 (2013) 965–1003.

[29] M. Franzese, A. Iuliano, Correlation analysis, in: P. Ranganathan (Ed.), Encycl. Bioinforma. Comput. Biol., Elsevier, 2018, pp. 706-721. https://doi.org/10.1016/B978-0-12-809633-8.20358-0.

[30] J.C.F. De Winter, S.D. Gosling, J. Potter, Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data, Psychological Methods 21 (2016) 273–290.

[31] E.J. Choi, Building thermal control based on real-time metabolic rate and clothing insulation of occupants by image data, Doctoral Dissertation, Chung-Ang Univ., 2023.

[32] J. Wang, J. Xu, X. Wang, Combination of hyperband and Bayesian optimization for hyperparameter optimization in deep learning, arXiv preprint arXiv:1801.01596. 10.48550/arXiv.1801.01596.

[33] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A. Boulesteix, Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 13 (2023) e1484, https://doi.org/10.1002/widm.1484.

[34] J. Bergstra, B. Komer, C. Eliasmith, D. Yamins, D.D. Cox, Hyperopt: a python library for model selection and hyperparameter optimization, Comput. Sci. Discov. 8 (2015) 014008, https://doi.org/10.1088/1749-4699/8/1/014008.

[35] J. Wu, X. Chen, H. Zhang, L. Xiong, H. Lei, S. Deng, Hyperparameter optimization for machine learning models based on Bayesian optimization, J. Electron. Sci. Technol. 17 (2019) 26–40, https://doi.org/10.11989/JEST.1674-862X.80904120.

[36] ANSI/ASHRAE Standard 62.1-2019, Ventilation for Acceptable Indoor Air Quality, ASHRAEASHRAE, Atlanta, 2019.

[37] W. Chan, S.-C. Lee, Y. Chen, B. Mak, K. Wong, C.-S. Chan, Z. Cheng, X. Guo, Indoor air quality in new hotels' guest rooms of the major world factory region, Int. J. Hosp. Manag. 28 (2009) 26–32, https://doi.org/10.1016/j.ijhm.2008.03.004.